

**Immagine e intento.
Viaggio nel potenziale abilitativo
delle IA generative**

**Image and Intent.
Journey into the Empowering Potential
of Generative AIs**

Lorenzo Ceccon¹, Matteo Cavaglia²

¹Politecnico di Milano
Dipartimento di Architettura e Studi Urbani
lorenzo.ceccon@polimi.it

²Politecnico di Milano
Dipartimento di Architettura, Ingegneria delle Costruzioni e Ambiente Costruito
matteo.cavaglia@polimi.it



fotografia
IA generative
immagini sintetiche
generazione di immagini
semiotica

photography
generative AI
image synthesis
image generation
semiotics

L'IA generativa ha suscitato notevole interesse nel mondo dell'arte e della fotografia. Interesse però non sempre di natura propositiva, in quanto la mancanza di regolamentazioni condivise pone numerosi rischi riguardo l'etico sfruttamento di questa innovativa tecnologia. Il presente lavoro è stato svolto all'interno dello scenario presentato, con l'obiettivo di verificare alcuni possibili paralleli tra le questioni e le critiche sviluppatesi attorno la nascente 'arte' creata con IA generativa e il quadro culturale proprio dell'arte fotografica. L'analisi ha portato a evidenziare alcune somiglianze, ma anche differenze, nel grado di autonomia esistente tra intento dell'artista e risultato. In quest'ottica è stata condotta un'analisi semiotica basata su sperimentazione empirica rispetto la verifica della riconoscibilità dell'intento 'fotografico' dell'artista che fa uso di AI all'interno del risultato visuale ottenuto, così come reinterpretato da algoritmi di classificazione delle immagini. I risultati paiono confermare la compatibilità dell'idea che anche l'IA generativa possa dotare l'artista digitale dei mezzi necessari a esercitare una forma di controllo all'interno del processo di generazione dell'immagine. Questa forma di controllo è ovviamente estremamente peculiare, ma in senso lato può essere ritenuta analoga alle modalità operative su cui la tecnica fotografica si è sviluppata. La possibilità di collegare intento e risultato, pare aprire delle interessanti prospettive di utilizzo della 'fotografia tramite IA', in particolare rispetto al ruolo dell'operatore umano coinvolto. La componente umana del processo diviene infatti 'pilota' della tecnologia, e viene qui considerata una forma di partecipazione fondamentale per 'umanizzare' il processo creativo così delineato.

Generative AI has sparked considerable interest in the world of art and photography. However, this interest is not always of a proactive nature, as the lack of shared regulations poses numerous risks regarding the ethical exploitation of this innovative technology. This work has been carried out within the presented scenario with the aim of verifying some possible parallels between the issues and criticisms that have emerged around the emerging 'art' created with generative AI and the cultural framework specific to photographic art. The analysis has highlighted some similarities, as well as differences, in the degree of autonomy existing between the artist's intent and the result. In this perspective, a semiotic analysis based on empirical experimentation was conducted to verify the recognizability of the 'photographic' intent of the artist using AI within the obtained visual result, as reinterpreted by image classification algorithms. The results seem to confirm the compatibility of the idea that generative AI can also provide the digital 'artist' with the means necessary to exert a form of control within the image generation process. This form of control is obviously extremely peculiar but can be broadly considered analogous to the operational methods on which photographic technique has developed. The possibility of connecting intent and result appears to open interesting prospects for the use of 'photography through AI,' particularly regarding the role of the involved human operator. The human component of the process becomes, in fact, the 'pilot' of technology and is considered here a fundamental form of participation to 'humanize' the creative process outlined.

Introduzione

La fotografia può essere sinteticamente definita come quell'arte resa possibile dall'uso dello strumento denominato 'macchina fotografica'. Questo legame vitale con uno strumento tecnologico contraddistingue la fotografia per la sua caratteristica di svilupparsi attraverso la spinta di due forze intrecciate, ma sempre distinguibili. Da un lato lo sviluppo culturale, che ha permesso dapprima il riconoscimento e infine il consenso, per un *medium* espressivo diverso da ogni altra arte, o tecnica, vista nella storia dell'uomo (Reilly, 1979; Teicher, 2016). Dall'altro lato, la forza dello sviluppo tecnologico, il quale ha permesso nel tempo di trasfigurare quell'insieme quasi alchemico di complesse procedure, in una pluralità di possibilità sempre più accessibili al vasto pubblico. Lo sviluppo tecnologico della fotografia è indubbiamente una storia di apertura indirizzata ad abilitare un pubblico sempre più ampio all'uso di questo *medium*.

Tuttavia, il progresso tecnologico fin qui citato ha anche portato a nuove sfide etiche legate alla manipolazione delle immagini e alla loro autenticità, in particolare, attraverso il passaggio dalle tecnologie analogiche a quelle digitali. Un cambio epocale, che definì addirittura la possibilità di fotografare con sempre più facilità non solo la realtà, ma persino l'irreale (Shapley, 2011; Shen et al., 2019).

Oggi, si può ritenere di essere alla soglia di un'altra transizione, innescata da una nuova tecnologia, che forse ancora più di prima, pone dinanzi a noi un eguale orizzonte di possibilità, quanto di responsabilità, su come comprendere e orientare questa rivoluzione in atto. La tecnologia in questione è l'Intelligenza Artificiale (IA) applicata alla generazione di immagini.

Il seguente contributo si pone in quest'ottica come un'esplorazione del presente orizzonte nel campo delle IA generative. Espo- nendo una metodologia di indagine, che viene qui considerata come un'applicazione di sperimentazioni di carattere maggiormente empirico che teorico, volte a raccogliere diverse suggestioni indirizzate nel regolare e formalizzare l'interazione tra 'creatore' e AI, senza risparmiare osservazioni di carattere pratico e tecnico, ma anche semiotico e talvolta etico.

Infatti, se già in passato il dibattito nell'ambito delle arti è stato influenzato dalla semiotica – e in particolare dal legame tra referente, significante e significato – le opportunità offerte dai sistemi di intelligenza artificiale generativa oggi disponibili sem-

brano offrire spunti di riflessione ulteriori, obbligando una rinnovata interrogazione su un possibile cambio di paradigma in ciò che l'autore dell'immagine si ripropone, e su ciò che costituisca il suo reale contributo di paternità, nonché le sue responsabilità. A livello più concettuale, ci pare che, nonostante l'IA non si possa considerare per nulla senziente, il suo utilizzo debba implicare l'applicazione del triangolo semiotico considerando l'IA come un soggetto che a sua volta collega significanti a significati, sebbene secondo una logica *'fuzzy'*, in modo non sempre univoco, o trasparente. Tale 'soggetto', per giunta, dipende non solo dalla piattaforma e dai parametri algoritmici – inclusi parametri di casualità – presenti nell'algoritmo, bensì soprattutto dai database (entità dinamiche, in realtà, e non statiche) che sono alla base dei collegamenti tra le immagini e il loro significato.

Il processo di comunicazione tra l'agente umano e l'IA è spesso basata su un binomio di tipo 'richiesta-risultato' (Witteveen & Andrews, 2022; The Ultimate Prompting Guide, 2022), e per questo motivo essa presenta nella sua fruizione l'implicita premessa che in ultima analisi la macchina sia in grado di eseguire quanto a essa richiesto. Al contrario però, *bias* di varia tipologia possono portare la macchina a deviare anche sostanzialmente dall'*input* iniziale o, meglio, da ciò che il creatore si aspetta quando utilizza un certo *prompt*. Per questo motivo, l'azione delle IA, pur in linea di principio operante in un'ottica deterministica, si sviluppa così attraverso un reticolo di variabili tanto complesso da sviluppare una sorta di propria 'sfocatura termodinamica', secondo cui l'utente può solo prevedere 'sfocatamente' l'esito delle proprie azioni, osservandone chiaramente l'aspetto solo a fatto compiuto (fig. 2).

Un processo creativo che include l'uso delle IA è per noi quindi non solo possibile, ma in ultima analisi reso proficuo proprio dalla partecipazione dell'attore umano, come 'pilota' di questa nuova tecnologia, in modo non dissimile alla fotografia del passato. Una fotografia che non solo arriva a fotografare l'irreale, ma persino ciò che è sconosciuto all'utente stesso.

In queste condizioni l'IA svolge un vero e proprio ruolo di 'agente abilitante', favorendo tramite la propria esistenza nuove 'tecniche', in modo simile alla 'macchina fotografica' stessa.

Il potere dell'abilitare ciò che prima non era possibile è però incredibilmente importante, e delicato. In questo frangente infatti tutti i professionisti coinvolti nello sviluppo di queste tecnologie sono chiamati a rispondere a forti interrogazioni di carattere

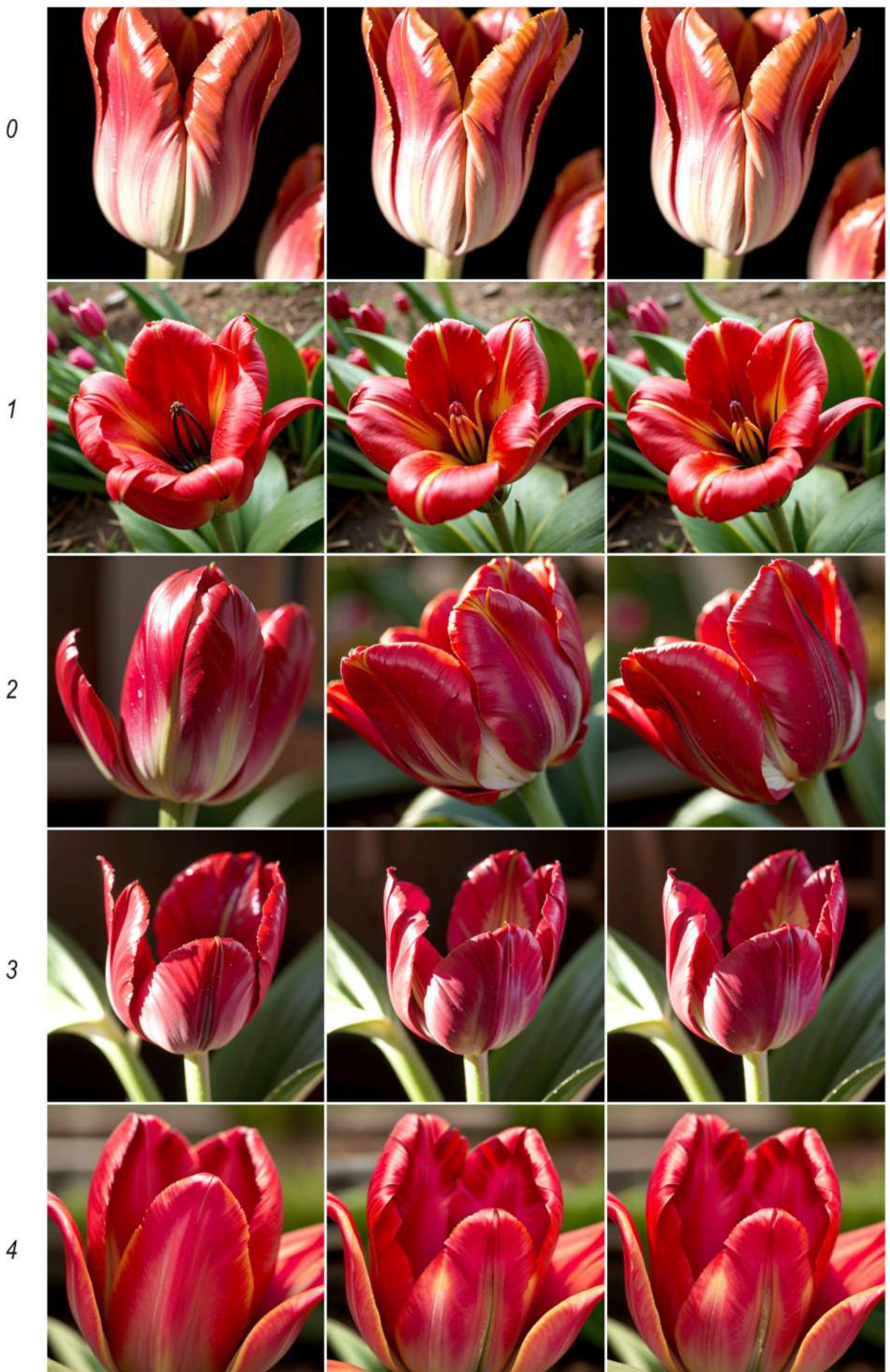
Fig. 1
Elaborazione grafica ad opera degli autori. L'immagine include contenuti derivati dall'uso di IA generativa. Immagini create utilizzando il *prompt* "A photo of a beautiful riverbank in an urban context" da Midjourney.

Fig. 2
Elaborazione grafica degli autori. L'immagine include contenuti derivati dall'uso di IA generativa. Immagini create utilizzando il *prompt* "A macro photo of a red tulip seen frontally at ground level" da Stable Diffusion. L'immagine illustra le variazioni nei risultati generati da uno stesso *prompt*, con aggiustamenti apportati a variabili specifiche. In verticale, i valori di *seed* variano in un intervallo compreso tra 0 e 4. In orizzontale, l'immagine mostra un confronto nell'utilizzo di *sampler* diversi.

EULER

HEUN

DPM++ 2M KARRAS



etico, in modo tale da evitare che in conclusione, non vengano ‘abilitate’ anche implementazioni dannose e nocive per i professionisti coinvolti, o la società stessa nel suo insieme. A questo riguardo, sono difatti innumerevoli gli allarmi rivolti non solo all’interno del dibattito specialistico, ma anche alla società nel suo insieme, con l’intento di denunciare i numerosi rischi che potrebbero essere determinati dall’uso improprio dell’Intelligenza Artificiale legato alla produzione di immagini (Gu et al., 2022; Westerlund, 2019).

La prospettiva temporale per valutare con completezza il reale impatto di queste tecnologie è forse però ancora troppo ristretta. È in ogni caso auspicabile che ulteriori studi forniscano maggiori evidenze per guidare un uso consapevole e positivo di questa tecnologia.

In questo contesto, la storia della tecnica fotografica offre interessanti parallelismi di straordinaria attualità. In particolare, vale la pena menzionare il saggio *Photography* (1855), pubblicato nel 1855, che è diventato noto per presentare in modo provocatorio alcune opinioni diffuse nel XIX secolo riguardo al mezzo della fotografia.

Si tratta di pensieri oggi giorno incomprensibili, ma curiosamente simili, per forma e contenuto, ad alcune osservazioni contemporanee nel dibattito riguardo l’uso artistico dell’Intelligenza Artificiale (McClard, 2022). Riportiamo ad esempio le seguenti frasi: “Ma anche questa applicazione è limitata alla mera riduzione e copia di opere precedentemente incise o disegnate; infatti, per quanto ingegnoso sia il processo o sorprendenti i risultati della fotografia, bisogna ricordare che quest’arte aspira solo a copiare, non può inventare” (*Photography*, 1855), o ancora, “Come ausiliaria dell’arte, la fotografia viene opportunamente in aiuto dell’artista in quegli studi preparatori sempre necessari nella produzione di un’opera d’arte” (*Photography*, 1855).

Il lavoro per sistematizzare questi sistemi sarà probabilmente arduo, ma l’esempio della fotografia può dimostrare come questa sia una scelta che presenta un potenziale che al momento forse ancora non può essere pienamente compreso.

Il funzionamento dei sistemi generativi di immagini

Qualsiasi tipologia di confronto intellettuale incentrato su una tecnologia complessa, necessita però come prima cosa, di

esplicitare, almeno parzialmente, i meccanismi di funzionamento di tale tecnologia. Seppur sia qui riconosciuto che il presente contesto risulti inadatto a un approfondito esame dei sistemi di Intelligenza Artificiale per la generazione automatica di immagini, si ritiene in ogni caso utile introdurre alcuni concetti essenziali per poter comprendere e delineare in modo sufficientemente chiaro limiti e possibilità di questi strumenti.

Infatti, ogni strumento può facilitare o, al contrario, ostacolare determinate azioni a causa delle sue caratteristiche intrinseche. Solo conoscendo tali caratteristiche, si può sviluppare una tecnica efficace per sfruttare appieno lo strumento scelto al fine di orientare l'uso anche in direzioni che sarebbero meno 'favorite' dal sistema stesso. Proseguendo il parallelismo con la macchina fotografica, così come una macchina fotografica può abilitare la riproduzione automatica dell'immagine di un contesto reale, allo stesso tempo essa impone dei limiti nell'esercizio di questa azione. Questi limiti possono essere di vario tipo, possono infatti appartenere genericamente alla tecnologia nel suo insieme, oppure essere propri dello specifico strumento in uso. La fotografia, come tecnologia nel suo insieme, basandosi sulla elaborazione delle informazioni luminose riflesse dalla scena da riprodurre, non può ad esempio essere adeguatamente implementata in situazioni in cui le informazioni luminose nell'ambiente sono insufficienti. Per questo motivo, la rappresentazione fotografica di una scena priva di luce, è intrinsecamente difficoltosa, se non del tutto impossibile. Tuttavia, l'utilizzo di illuminatori, di pellicole o sensori ad alta sensibilità, di lenti a basso *f-stop* e di tecniche specifiche per le condizioni di scarsa luminosità consentono di ottenere ottimi risultati in condizioni che a priori potrebbero apparire come proibitive: pensiamo alle scene girate a lume di candela da Stanley Kubrick in *Barry Lindon*! Addirittura, pensando alle sperimentazioni di Man Ray con una porzione del processo fotografico, si sono ottenuti esiti artistici che davvero poco avevano a che fare con la 'copia' della realtà.

Un discorso analogo, seppur diverso negli specifici 'termini', è del tutto applicabile al mondo della generazione generativa delle immagini. Per prima cosa, è necessario stabilire il meccanismo di base che permette ai sistemi di intelligenza artificiale individuati di operare. Mentre la macchina fotografica si basa sull'elaborazione della luce, questi sistemi digitali operano nel proprio livello più

essenziale come elaboratori di dati. L'elaborazione di questi dati avviene attraverso l'implementazione di specifici modelli generativi. Un modello generativo è un tipo di modello statistico, che permette di eseguire quello che viene spesso definito “apprendimento automatico” (Dean, 2023), con il fine sviluppare dati sintetici sulla base di riferimenti simili catalogati in specifici ‘set di dati di addestramento’. Con dato ‘sintetico’, si intende così un tipo di dato analogo a quelli forniti come esempio, ma non completamente uguale a nessuno di essi, e perciò potenzialmente ‘originale’, seppur in ogni caso collegato all’insieme di riferimento (Alam, 2022).

Nel caso dei sistemi specializzati nella generazione di immagini, il dato sintetico generato dal sistema, altro non sarà che l’immagine stessa richiesta dall’utente tramite il *prompt* di richiesta. E in particolare, è proprio il *prompt* a fornire la guida per utilizzare il vasto database alla base del modello all’immagine finale ottenibile dall’azione del modello generativo stesso.

Nel corso degli anni sono state studiate e sviluppate diverse tipologie di modelli generativi, ma gli strumenti di generazione automatica di immagini di maggior successo rilasciati (p. es. *DALL·E 2*, s.d., *Stable Diffusion*, s.d., e *Midjourney*, s.d.), sfruttano alla loro base un modello generativo di diffusione (*diffusion model*). I modelli di diffusione affiancano altri modelli analoghi (p. es. *GAN*, *VAE*, e modelli *Flow-based* (Karpathy et al., 2016)), e secondo diverse fonti rappresentano al momento lo stato dell’arte rispetto applicazioni di generazione di immagini (Dhariwal & Nichol, 2021; Inhal, 2022).

Il modello generativo di diffusione verte sul concetto di asse temporale, e prende ispirazione dal fenomeno di diffusione che è possibile esaminare negli ambiti della chimica e della fisica. In particolare, il fenomeno di diffusione identifica un qualunque spostamento di una sostanza tra una regione di alta concentrazione, a una di bassa concentrazione. Un esempio comune del fenomeno fisico di diffusione nella vita di tutti i giorni è la diffusione dell’odore. Quando viene spruzzato un profumo in una stanza, l’odore si diffonde gradualmente nell’aria e si può infine percepire in tutto l’ambiente. Questo semplice fenomeno identifica un evento temporalmente ben definito in graduali passaggi (il profumo si diffonde gradualmente, e non in un istante), che inoltre è da considerarsi irreversibile, poiché il profumo una volta diffuso non può ritornare spontaneamente nel flacone che lo conteneva.

Questa limitazione però non è applicabile all'ambiente digitale, dove invece potrebbe essere potenzialmente possibile tenere memoria di tutti gli istanti del processo, per poi avvolgerli liberamente avanti e indietro, come in un video.

La libera manipolazione della linea temporale rispetto a due eventi tra loro collegati è il campo d'azione dei modelli di diffusione, i quali vengono addestrati per estrapolare i pattern di causalità necessari a invertire un processo grafico specifico, cioè l'aggiunta di rumore a un'immagine (Muppalla & Hendryx, 2022). I modelli di diffusione sono quindi addestrati sottoponendo come 'set di dati di addestramento', immagini gradualmente sottoposte a un incremento di rumore grafico, fino alla completa distruzione dell'immagine originale, sostituita così dal solo rumore. Il rumore qui definito è ovviamente inteso nella sua declinazione digitale, cioè come la variazione imprevista e 'casuale', dei dati che compongono un'immagine. In sostanza, è come se il modello potesse invertire la naturale tendenza entropica dei sistemi fisici invertendo non solo il verso dell'asse temporale, ma anche quello dei nessi causali.

Un modello di diffusione con un addestramento sufficientemente vario può così implementare la capacità di operare l'operazione inversa, rimuovendo il rumore da un file grafico contenente informazioni grafiche casuali, con l'obiettivo di ricostruire l'immagine associata al *prompt* richiesto dall'utente. Il fatto che l'immagine richiesta non 'esista', non presuppone qui una limitazione, in quanto i dati di addestramento sono in grado di fornire un riferimento per dirigere questo processo di 'inversione temporale', verso un punto di partenza (l'immagine sintetica) plausibile (fig. 3) (Florian, 2023; Wankhede, 2023), in cui, dato il sistema complesso si possa comunque utilizzare, se non dei veri e propri nessi causali, quantomeno delle correlazioni statistiche basate sulle occorrenze del database alla base del modello.

Il modo in cui un *prompt* può condizionare e guidare la generazione del contenuto richiesto, piuttosto che un altro, può essere implementato in due modalità differenti: *Contrastive language image pre-training (CLIP)* (Nichol et al., 2021) e *Classifier-free guidance (CFG)* (Ho & Salimans, 2022). Le metodologie non sono tra loro equivalenti, ma in ogni caso nella loro esecuzione risultano ancora dipendenti dal 'set di dati di addestramento'. Questa dipendenza presenta un rischio intrinseco, che Zemel et.al. hanno sinteticamente descritto in questo modo: "I sistemi

addestrati a prendere decisioni basate su dati storici ereditano naturalmente i pregiudizi del passato” (Zemel et al., 2013). In questo contesto, il pregiudizio, o *bias* dall’inglese, è riferito alla potenziale e sistematica incorretta conclusione a cui un sistema sarebbe portato a giungere, rispetto a dati di riferimento imparziali, incompleti o a loro volta non corretti. In definitiva, come evidenziato da Barocas e Selbst (2016), un algoritmo è essenzialmente buono tanto quanto i dati con cui lavora.

Indagine sulla fotografia artificiale.

Il *bias* e la comunicazione con un soggetto non umano

È stato finora mostrato che un sistema generativo per la produzione di immagini può, in qualche modo, deviare dal suo obiettivo, rischiando di disorientare un utente umano che non riesca a interagire con il processo generativo in maniera adeguatamente consapevole e controllata, rischiando di trovarsi di fronte ad immagini che non corrispondono alle aspettative del creatore, e/o che mostrino un *bias*. Ci si chiede, pertanto, per quale motivo un sistema generativo rappresenterebbe comunque uno strumento fortemente attrattivo, e in grado di contribuire a generare un notevole impatto all’interno dei più disparati processi di lavoro nonostante tali rischi?

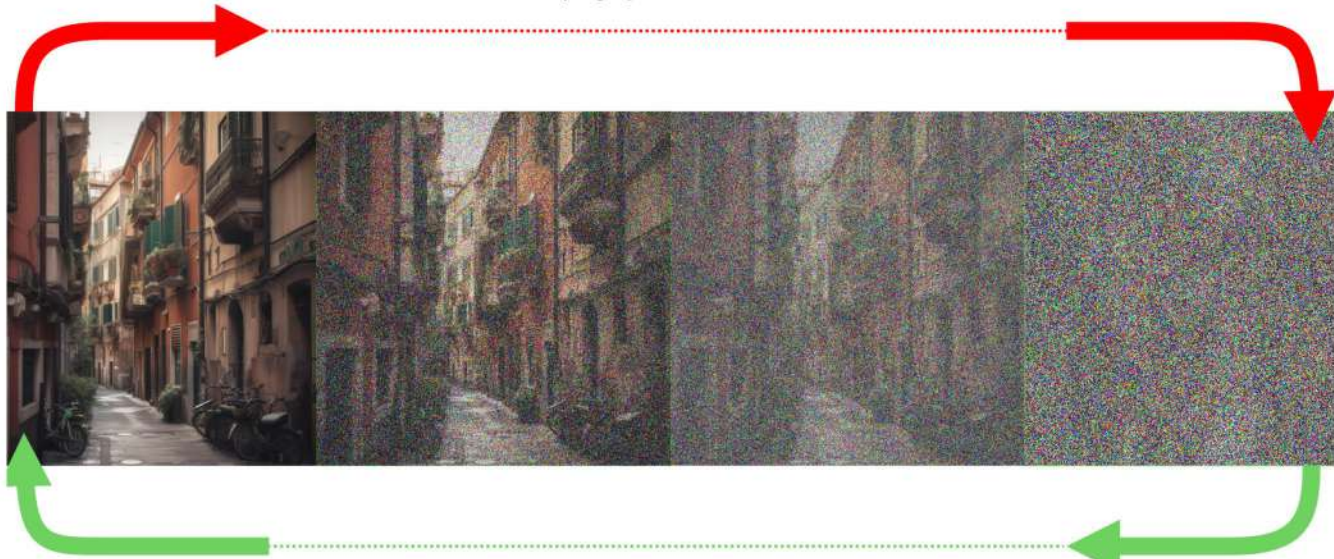
A questo interrogativo si può cercare di rispondere se esaminiamo il funzionamento di questi strumenti di IA e il loro rapporto con l’artista digitale, da un lato, e con i dati su cui vengono educati, dall’altro. Se è infatti vero che ogni strumento che usiamo tenderà a portarci a esiti formali che in parte dipendono da esso – un affresco e una scultura di un blocco di marmo porteranno ad esiti differenti anche se per mano del medesimo autore – nel caso dell’arte fotografica e, ancor più, della “fotografia” con IA l’autore si trova di fronte a mezzi i cui risultati finali non sono sempre del tutto prevedibili. Da questo punto di vista, pare che si torni all’origine della fotografia tradizionale, quando la profondità di campo poteva essere solo stimata in modo approssimativo, o alla latitudine di posa, così dipendente dal substrato fisico della pellicola specifica (oggi dal sensore specifico).

Il tema del grado e del tipo di ‘controllo’ dell’artista sull’*output* finale è già rilevante in relazione al *medium* fotografico tradizionale, ma anche nell’ambito della produzione di

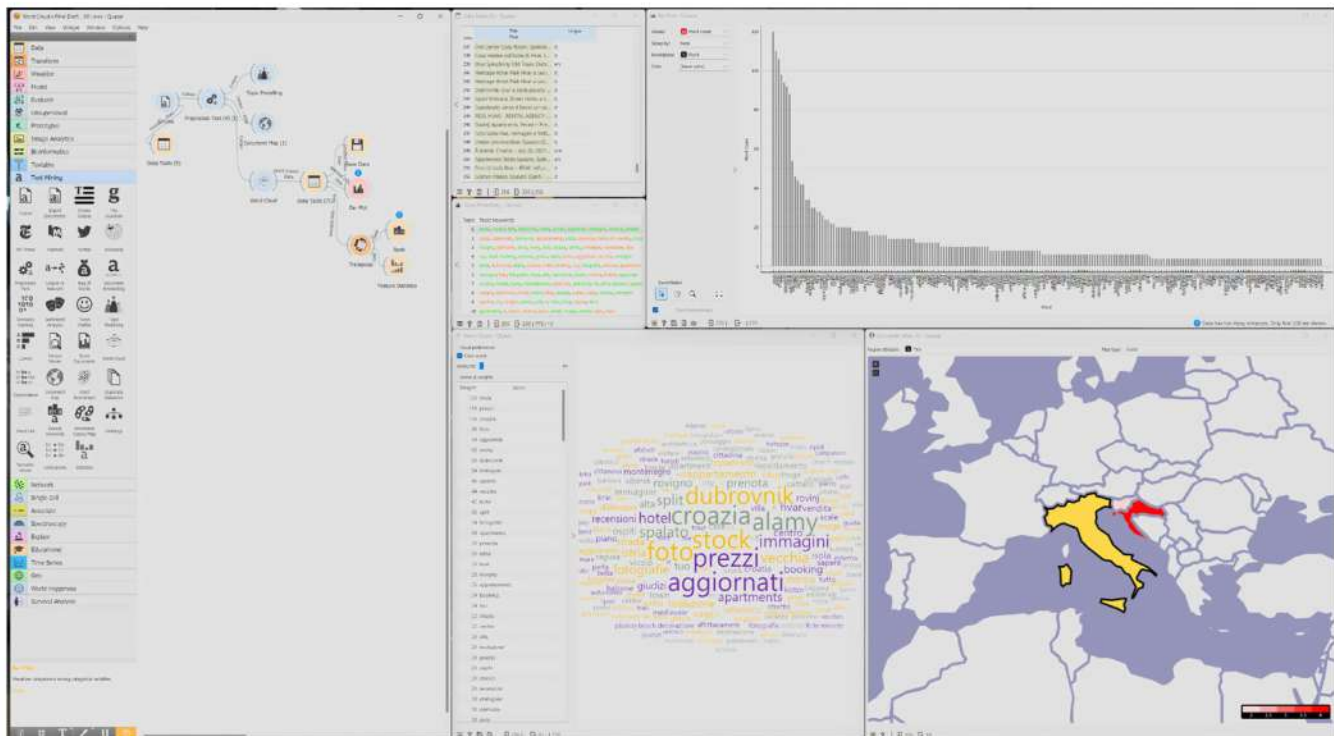
Fig. 3
Rappresentazione semplificata del processo di lavoro fondamentale di un modello a diffusione. L’immagine utilizzata per dettagliare l’esempio qui riportato è anch’essa frutto dell’applicazione di un sistema di generazione di immagini basato su intelligenza artificiale (*Midjourney*, s.d.)

Fig. 4
Elaborazione grafica ad opera degli Autori. L’immagine mostra un algoritmo in *Orange Data Mining* – linguaggio di Programmazione Visuale – che analizza l’output generato dall’IA in base ai *prompt*, in questo caso “*A photo of a beautiful Croatian urban context*”.

L'immagine originale viene lentamente trasformata in un campo grafico di dati casuali



Il modello generativo assimila in questo modo i pattern di editing per generare un'immagine a partire da un campo grafico di dati casuali



arte più in generale: pensiamo a tutte quelle espressioni artistiche che si focalizzano più sul processo che sul risultato specifico, che sia l'*action painting* di Jackson Pollock o sia la musica aleatoria di John Cage, Bruno Maderna e altri.

Ciò non toglie che nel caso dell'IA non si abbia mera aleatorietà, bensì anche un bagaglio di 'conoscenze' pregresse. I dati su cui i sistemi vengono addestrati – non tanto ci interrogano sul rapporto tra l'intento e il processo artistico, da un lato, e il risultato, dall'altro – bensì ci impongono anche di considerare l'IA una sorta di soggetto semi-autonomo, che contiene in sé elementi culturali derivanti dai dati di addestramento. La 'realtà cognitiva' dell'IA, infatti, porta ad attribuire al *prompt* di interrogazione l'ambito di significato che scaturisce dalla messe di dati del modello in modo non deterministico, ma comunque fortemente correlato in termini statistici: l'associazione più ricorrente tra immagini e *tag*, ad esempio, definirà l'area semantica di quel *tag*, con una variazione casuale dovuta allo specifico *seed* (seme di casualità) e secondo a una logica *fuzzy*, come già richiamata.

Pertanto, dato che l'IA si basa su dati che interpreta in modo che potremmo definire statistico – secondo una sorta di implicita stima dell'*id quod plerumque accidit*, cioè di ciò che si verifica statisticamente in modo più ricorrente, unita a dei *seed* di casualità – i suoi risultati tendono a portare in sé tali molteplici correlazioni statistiche, il che viene spesso definito *bias*, con connotazione spesso negativa, come se ci si trovasse di fronte a un errore che comporta il perpetuarsi di pregiudizi.

Ma è proprio nel rapporto tra intento dell'artista digitale, *prompt* e orizzonte semantico dell'IA che sta uno dei maggiori temi che vale la pena analizzare per meglio valutare la collocazione degli strumenti di IA nell'ambito della creazione di immagini 'fotografiche'.

Bias e triangolo semiotico.

Se, infatti, da un lato tale *bias* viene spesso giustamente aborrito, in particolare come veicolo di convinzioni o associazioni non politicamente corrette, tramite l'interpretazione fallace di temi estremamente sensibili come la rappresentazione stereotipata di etnie o minoranze, dall'altro lato dobbiamo riflettere più

a fondo su come noi stessi umani utilizziamo il linguaggio per verificare se il bias sia un elemento evitabile o ineliminabile nella comunicazione verbale.

In un certo senso, è infatti possibile sostenere che i bias siano una caratteristica ineluttabile di una qualsiasi conversazione tra soggetti autonomi. In un qualsiasi dialogo, infatti, è pienamente possibile e probabile che non vi sia piena concordanza tra il significato attribuito alle parole utilizzate nella conversazione da parte di chi le usa, e il significato da esse recapitato a parte ricevente, sia per le evidenti aree di incertezza semantica dell'espressione linguistica, che per il fatto stesso che il significato – anche ad essere certi di una perfetta corrispondenza semantica tra i mondi dei due soggetti – dipenderà anche dal 'mondo mentale', si potrebbe dire dalla *Weltanschauung*, dei due soggetti. L'impatto di questa incomprendimento però, non sempre è tale da compromettere il senso della conversazione stessa; per questo motivo, mentre da un lato è giusto affrontare e chiarire ogni incomprendimento che porti a fraintendere palesemente e grossolanamente il messaggio desiderato, molto più spesso è comunque possibile comunicare in modo efficace pur trascurando malintesi di ordine inferiore, e anzi considerandoli come una ovvia realtà nelle nostre comunicazioni verbali e non. Anzi, a volte sono proprio questi 'malintesi' che potremmo definire di 'bassa magnitudo', a determinare l'incremento di valore dell'incontro, poiché aprono la strada verso la comprensione di un punto di vista diverso dal nostro, posto che riusciamo a comprenderne la portata, oltre che a coglierne l'esistenza.

Volendo semplificare, potremmo così suddividere i *bias* in due categorie: i *bias* ad alto impatto, che rappresentano errori in grado di compromettere in modo palese e fondamentale la comunicazione, e *bias* di minore impatto, che invece identificano una sorta di 'sfumatura' nell'interpretazione di determinate idee o concetti, pur mantenendone univoca l'interpretazione di fondo.

In tema di *bias* ad alto impatto, essi comportano infatti la manifestazione di fondamentali errori in un dato sistema, che, come tali, devono essere gestiti. Citiamo ad esempio la ricerca di Joy Buolamwini e Timni Gebru., che ha evidenziato come i sistemi di riconoscimento facciale possano essere influenzati dal *bias* di genere e razziale, portando a risultati imprecisi per le donne e le persone di colore (Buolamwini & Gebru, 2018). In altre parole, in questi casi il sistema non riconosce correttamente persone che ricadono in una

categoria sottorappresentata nei dataset di addestramento. In questi casi è opportuno considerare attentamente su quali dati e con quali modalità si addestrino le IA, ma che ci troviamo di fronte a un *bias* ad alto impatto è dimostrato dal fatto che abbiamo degli ‘errori’ di riconoscimento vero e proprio. Il sistema, insomma, non è efficace nel compito per cui è stato concepito.

In sostanza, parlando di ‘fondamentali errori’, si riscontra in questo ambito ciò che in ambito giuridico è chiamato ‘*aliud pro alio*’, e che caratterizza la disciplina che collega la descrizione dell’oggetto del contratto di compravendita e l’oggetto specifico venduto. Infatti, per *aliud pro alio* si intende “quando viene consegnato un bene completamente diverso da quello pattuito” (Gazzoni, 1994, p. 1077). La difficoltà nel quantificare o nel qualificare cosa si intenda per ‘completamente diverso’, e quindi quando ci si trovi di fronte a un caso di *aliud pro alio*, deriva dal fatto che, come si è visto, ci troviamo in un caso di semantica *fuzzy*, cioè aree semantiche che non hanno un confine netto, ma una sorta di limes, di area semantica incerta a cavallo tra essi. L’*aliud pro alio* e, nel nostro caso, il *bias* ad alto impatto non avviene quando parola e suo significato concreto – immagine nella fotografia IA, e bene oggetto di vendita nella compravendita – ricadano nelle aree semantiche dubbie, aree in cui appunto avremmo un *bias* di bassa intensità, bensì solo ove ci si trovi in un campo certamente differente da quello usualmente inteso come corrispondente a tale parola.

Si noti ancora che, nell’ambito della fotografia generativa con IA, le parole relative a sensazioni, emozioni e concetti astratti tendono ad avere un’area semantica più *fuzzy* di quelle che indicano cose, e quindi sono più prone a *bias*: se scrivo “albero” o “casa” è più probabile che non si avverta tale *bias* rispetto all’aggettivazione emotiva legata alle cose. Se scrivessi “albero allegro” o “casa accogliente” non è detto che riuscirei a tracciare a ritroso quale caratteristica l’*output* visuale stia a rappresentare. Sfumature di linguaggio, che aumentano le aree di incerta attribuzione semantica, insieme a diverse coordinate concettuali – una casa “accogliente” può appartenere a gruppi diversi secondo altre dimensioni semantiche, come ad esempio stile, epoca, stato di manutenzione ecc. – aumentano le chance di *bias* a basso impatto senza necessariamente aumentare quelle di *bias* ad alto impatto.

Quanto al *bias* a basso impatto, infatti, esso può fortemente influenzare la creazione di arte attraverso l’IA. Ad esempio, un

algoritmo di IA addestrato su un insieme di dati che include principalmente opere d'arte raffiguranti soggetti di genere maschile, con tratti occidentali e carnagione bianca, potrebbe generare contenuti che riflettono questi stessi *bias*, cioè, mostrando con maggiore frequenza questo tipo di soggetti. Questo è un problema che è stato discusso da Elgammal et al. (2017), che hanno esaminato come l'IA può essere utilizzata per creare nuove opere d'arte. In questo esempio, appunto, se da un lato un *prompt* relativo a un "uomo adulto" probabilmente non commetterebbe errori di *bias* ad alto impatto – mostrando correttamente un soggetto umano di sesso maschile e di età adulta – quasi certamente mostrerebbe un uomo dai tratti Europei e con i segni di maturità tipici di tale gruppo umano: barba, acconciatura ecc. In altri termini, non avremmo un *aliud pro alio*, bensì un sottoinsieme di ciò che potremmo intendere con l'espressione "uomo adulto" che non è errato in sé, ma risulta *biased* nella statistica rappresentazione di ciò che a scala globale "uomo adulto" potrebbe significare, rappresentandone solo un ridotto sottoinsieme.

Sempre in relazione al bias, in particolare a quello di minore impatto, notiamo che esso pare ineliminabile elemento di ogni sistema semiotico. Il linguaggio, come sistema di simboli e regole, è costituito da due elementi chiave: il significante e il significato. Il significante è la forma fisica del simbolo, mentre il significato è l'idea o il concetto che il simbolo rappresenta. Questo legame tra significante e significato è stato ampiamente studiato nella linguistica, in particolare da Ferdinand de Saussure (1916), che lo ha definito come "arbitrario e convenzionale". Inoltre, Charles Sanders Peirce (1954) ha introdotto il concetto del triangolo semiotico, che aggiunge un terzo elemento a questa relazione: l'interpretante. L'interpretante è la comprensione o l'interpretazione del significato da parte del ricevente. Questo triangolo semiotico rappresenta la complessità della comunicazione linguistica e sottolinea l'importanza dell'interpretazione nel processo di comunicazione.

Nel contesto dell'Intelligenza Artificiale (IA), il *bias* può quindi essere visto come un fenomeno inevitabile, influenzato da vari fattori, tra cui il modo in cui i dati sono raccolti, elaborati e interpretati. Non si tratta necessariamente di un 'errore', ma piuttosto di un riflesso di quell'interpretante che in questo caso altro non è che l'interfaccia statistica tra il significante e il significato, che può essere influenzata da basi di dati poco rappresentative o 'poco

corrette', cioè poco allineate con i nostri desiderata e le nostre convinzioni politiche, e quindi una sorta di scontro tra *Weltanschauung* diverse tra il soggetto umano che esprime il *prompt* e il soggetto IA che lo interpreta.

Un esempio di come il legame tra significante e significato possa influenzare il bias nell'IA è il *Natural Language Processing* (NLP), un campo dell'IA che si occupa dell'interpretazione e della generazione del linguaggio naturale. Nell'NLP, il significante è rappresentato dai dati di *input*, come le parole o le frasi, mentre il significato è rappresentato dall'interpretazione di questi dati da parte dell'IA. Se l'IA non è in grado di comprendere 'correttamente' il legame tra significante e significato, può generare risultati 'distorti'. Un esempio di questo tipo di bias può essere trovato in uno studio del 2021 (Bolukbasi et al. 2021), in cui è stato dimostrato che gli algoritmi di NLP possono mostrare un *bias* di genere, attribuendo stereotipi di genere a determinate professioni. Questo *bias* può essere attribuito a un'interpretazione 'errata' del legame tra significante e significato, in cui l'IA associa determinati significati (stereotipi di genere) a determinati significanti (professioni). Ad esempio, se chiedessimo di raffigurare un manager, l'output probabilmente consisterebbe in un uomo caucasico di mezza età, vestito all'occidentale. Questo esempio evidenzia come il *bias* nell'IA possa essere visto come un'interfaccia statistica 'interpretante' tra il significante e il significato, piuttosto che un 'errore', e spesso è d'altronde un mero specchio di una distribuzione di 'stereotipi' che effettivamente sono presenti non solo nei dati di addestramento, ma anche nell'*humus* culturale, anche 'non detto', all'interno della società.

Posto che il *bias* non possa essere definito come un semplice errore che si dovrebbe tendere ad eliminare, bensì come un aspetto necessariamente intrinseco al linguaggio stesso – come disallineamento tra significante e significato rispetto a soggetti/interpretanti diversi – diventa dunque fondamentale immaginare un modo per 'correggere' tale disallineamento per poter mantenere il controllo sulla generazione (anche) artistica tramite IA. Anzi, come vedremo, più che di 'correzione' si tratta di 'traduzione' tra interpretanti differenti.

A tal proposito, nell'articolo *Recipient design in human communication: simple heuristics or perspective taking?* (Blokpoel et al., 2012), gli autori discutono il concetto di *recipient design*, che si riferisce all'adattamento del comportamento comunicativo in

base al destinatario, un concetto strettamente legato al triangolo semiotico. Gli autori suggeriscono che l'IA potrebbe utilizzare meccanismi simili per adattare il suo comportamento comunicativo. Ad esempio, un sistema di IA potrebbe utilizzare euristiche semplici o prendere in considerazione la prospettiva del destinatario (ad esempio pesando il proprio feedback rispetto agli *input* dell'utente) per adattare la sua comunicazione. Ma è vero altresì anche il contrario: noto il *bias*, l'artista potrà adattare i propri *prompt* di richiesta al fine di allineare i significati, anche utilizzando termini che a suo giudizio non sarebbero corretti, e viceversa. Ad esempio, se scoprissi che un "bel paesaggio" viene interpretato come "un bel paesaggio naturale", ecco che – se voglio rappresentare un bel paesaggio che comprenda anche opere antropiche – potrò/dovrò aggiungere qualche indicazione al riguardo, operando essenzialmente una modifica della richiesta iniziale, cioè della comunicazione. Di fatto questa ultima soluzione non è poi molto diversa di quanto si tende a fare tra esseri umani: ad esempio, probabilmente non utilizzerò l'espressione "maggioranza Bulgara" in senso dispregiativo con chi fosse Bulgaro o nostalgico del regime cui l'espressione si riferisce, né utilizzerò in modo dispregiativo l'espressione "fascista" con chi sia a sua volta nostalgico di quel regime.

Conversare con l'IA, la rappresentazione del *Genius loci* nell'epoca dell'Intelligenza Artificiale

In linea con l'analisi sopra sviluppata, si sono state eseguite diverse sperimentazioni di carattere pratico, relative alle modalità di comunicazione e generazione di contenuti mediante la costruzione di *prompt* specifici, operando essenzialmente all'interno del dominio definito come *prompt engineering*, o *ingegneria dei prompt* (White, 2023). L'ingegneria dei *prompt* è un termine di recente formazione, utilizzato per identificare l'ambito tecnologico che studia l'ottimizzazione della comunicazione e utilizzo, dei modelli di linguaggio basati sull'intelligenza artificiale. Come molti altri sistemi informatici, anche i servizi di generazione automatica di immagini basati su IA, possono essere strutturarsi su due diversi livelli di accesso e gestione, il livello visibile all'utente e con cui esso può interagire con il sistema, ed il livello 'profondo' che invece si interfaccia con i componenti software

che permettono l'effettivo funzionamento del sistema stesso, rispettivamente chiamati *front-end* e *back-end*.

In particolare, il front-end dei sistemi di generazione automatica tramite IA basati sull'uso del linguaggio naturale, e non essendo quindi più 'limitati' sulla sola manipolazione di una ben definita interfaccia utente (UI, da *User Interface*), ma su modalità di 'apparente' conversazione libera basate sui *prompt*, fa sì che le strategie di utilizzo di questi ambienti non siano necessariamente focalizzate su un concetto di utilizzo 'procedurale' (cioè sulla costruzione di sequenze di operazioni logicamente ordinate tra loro), ma al contrario su un modello statistico 'semantico', ovvero basato sull'elaborazione verbale (in forma di testi) di concetti, con particolare rilevanza sull'uso 'comune' del linguaggio e delle parole (White, 2023).

Il binomio tra *frontend* e *backend*, non fa altro che identificare all'interno dell'ambito informatico, quella normale dualità insita in qualunque tecnologia umana, ovvero la distinzione tra coloro che 'usano uno strumento', e coloro che 'creano quello strumento'. Questa distinzione è particolarmente rilevante all'interno della ricerca inerente tecnologie abilitative con impatto nei settori dell'arte e della produzione creativa, in quanto seppur questi due gruppi possano spesso sovrapporsi, possiedono in un'ottica generale obiettivi marcatamente distinti. Obiettivi che, rinforzando le precedenti classificazioni, è possibile riassumere genericamente con l'obiettivo di 'sfruttare uno strumento per creare un contenuto di valore', da un lato, e 'definire uno strumento di valore per abilitare la creazione di contenuti' dall'altro lato (Laaziri et al., 2019). La distinzione ovviamente non è assoluta, ma è interessante da considerare per comprendere la complessità della partecipazione dell'utente finale di una certa tecnologia all'interno dello sviluppo della tecnologia stessa.

È talvolta rilevante ricordare come l'utente finale di una tecnologia non sempre necessiti di una comprensione profonda del suo funzionamento per sperimentarne l'uso e trarne valore. Prendiamo ad esempio la fotografia: dove sono moltissime le persone che utilizzano macchine fotografiche o smartphone per catturare immagini, senza necessariamente comprendere i più sofisticati meccanismi sottostanti di ottica, sensoristica e algoritmi di elaborazione delle immagini. L'obiettivo dell'utente finale è infatti sperimentare l'uso dell'oggetto, esplorare le sue potenzialità e creare qualcosa di unico attraverso il filtro dell'interfaccia utente. Questa sperimentazione, guidata dalla curiosità

e dalla creatività, può spesso portare a scoperte e intuizioni di grande interesse anche per gli esperti che lavorano con la tecnologia a un livello più profondo. In questo senso, ogni utente diventa un innovatore, un esploratore del possibile, contribuendo a plasmare il futuro della tecnologia stessa. Proprio per questo motivo, una tecnologia come l'intelligenza artificiale, dove l'interfaccia utente è di primo impatto ridotta ad una semplice "conversazione", offre un campo di osservazione particolarmente interessante, poiché gli utenti interagiscono con l'IA attraverso un processo di lavoro particolarmente fluido.

Tuttavia, come nel campo della fotografia la conoscenza delle caratteristiche ottiche di una certa lente – in termini di *bokeh*, *flair*, diffrazione ad alti diaframmi, etc – nonché del sensore o della pellicola specifici sono considerati una necessità da parte dei fotografi professionisti, che uniscono alla sperimentazione empirica anche queste nozioni più approfondite per controllare la propria espressione artistica, così anche nell'utilizzo dell'IA e del *promptengineering* è bene essere consapevoli sia della modalità di funzionamento di tali algoritmi in generale, sia dello specifico algoritmo e base dati, in modo da sperimentare con maggior consapevolezza. Insomma, senza voler escludere a priori che una sperimentazione istintiva possa portare a risultati interessanti, normalmente – che sia in abito fotografico tradizionale o in quello della creazione di immagini tramite prompt engineering – un certo controllo a priori (cioè non tramite mera selezione a posteriori di output generati a caso) sul risultato ottenuto dall'artista pare necessario anche per consentire quegli aspetti di riconoscibilità e coerenza nell'intento e nella forma dell'espressione artistica che può farci considerare l'artista tramite *AI-prompting* un artista e non un *editor*, né un puro fruitore di creatività altrui.

La maggior differenza con la fotografia tradizionale si ravvisa tuttavia nel fatto che, mentre in quel caso l'artista ottiene immagini per raggiungere i risultati emotivi desiderati immaginando quale tipo di immagine meglio possa evocare tali emozioni, nella fotografia con IA l'artista digitale potrà utilizzare direttamente all'interno del prompt parole dal significato emotivo – come "bello", "commovente", "tragico", ecc. – che a loro volta evocheranno nell'interpretazione dell'IA delle immagini che spazieranno in un range *fuzzy* di immagini cui tali parole sono connesse. Occorre dunque che l'artista IA sviluppi una comprensione il più possibile approfondita di tale implicita

‘traduzione’ semantica, in modo da poter, ove lo voglia, meglio controllare l’immagine ottenuta.

Per concludere, riportiamo ora alcune osservazioni estratte da una limitata sperimentazione eseguita nell’ambito dello sviluppo di immagini generative tramite il sistema di *Midjourney* (s.d.) incentrata sulla sintesi di immagini di scenari urbani.

È importante notare come la generica rappresentazione di un determinato contesto reale sia un compito affascinante, poiché oltre l’apparente semplicità cela invece un notevole grado di complessità. Complessità derivata principalmente dal confronto con l’evocativo concetto di *genius loci* proprio del luogo scelto per la rappresentazione (Bravo, 2010). L’immaginare un luogo in modo coerente implica infatti l’atto di mettere insieme un ricco mosaico di informazioni che contribuiscono a definire l’identità unica ad essa associata. Il *genius loci* rappresenta quell’elemento indefinibile e sottile che conferisce a un luogo la sua anima, quello che potremmo empiricamente definire come il suo “carattere distintivo” (Bravo, 2010). Misurarsi con il *genius loci* richiede una sensibilità particolare per cogliere le sfumature e i dettagli che rendono quel luogo unico. È solo attraverso questa comprensione profonda che possiamo creare spazi architettonici che risuonano con l’essenza del luogo, offrendo esperienze significative e connessioni autentiche, o almeno sufficientemente raffinate da imprimere all’interno della rappresentazione creata un significato con la potenzialità di stabilire un legame di riconoscibilità con l’osservatore (figg. 4-5).

Sulla stessa linea, chiedere a un sistema generativo di elaborare un contesto urbano ‘piacevole’ diventa il mezzo per sviluppare un’indagine sull’efficacia del sistema di Intelligenza Artificiale nell’interpretare e collegare in modo coerente le informazioni necessarie a eseguire tale compito non banale.

I risultati ottenuti da tali interrogazioni hanno evidenziato come *Midjourney*, nella sua configurazione attuale (versione 5.0), possa abilitare la generazione di immagini sintetiche con un grado di riconoscibilità piuttosto elevato. La verifica di questa conclusione è stata ottenuta, come anticipato, non solo attraverso la mera osservazione e discernimento umano, ma anche mediante un secondo sistema basato sull’intelligenza artificiale, utilizzato per valutare ciascuna immagine urbana elaborata. Sfruttando il motore di ricerca per immagini integrato in *Google*, è stato infatti possibile condurre ricerche web mirate, al fine di individuare ogni immagine creata all’interno dei

Fig. 5
Elaborazione grafica ad opera degli Autori. L’immagine include contenuti derivati dall’uso di IA generativa. Immagini create utilizzando il prompt “A photo of a common urban street located in Italy” da *Stable Diffusion* (s.d.). In questa elaborazione è stato mantenuto lo stesso valore di *seed* alternando però il contesto della scena ed il contesto geografico di riferimento. Ciò ha permesso di mantenere costanti numerosi parametri legati alla composizione delle immagini, ricercando invece l’interpretazione della macchina rispetto la variazione del *genius loci* proprio di ogni luogo.

URBAN STREET

URBAN SQUARE

SCENIC PLACE

ITALY



GERMANY



CHINA



SWEDEN



BRAZIL



domini web interrogabili dal motore di ricerca. Dato il carattere sintetico delle immagini, è ovviamente impossibile ottenere una corrispondenza completa con un'immagine esistente. Tuttavia, nel tentativo di soddisfare la richiesta data, il sistema effettuerà in ogni caso l'estrazione diversi risultati pertinenti in base ai criteri di 'similitudine' e 'somiglianza' (fig. 6).

I risultati hanno evidenziato come ogni immagine potesse essere collegata al web per ottenere non solo ad aree geografiche perfettamente coerenti con le richieste corrispondenti alla generazione delle immagini sintetiche (ad esempio, un'immagine generata dalla richiesta di rappresentare un contesto urbano italiano piacevole è stata correttamente collegata a fonti web relative a territori italiani), ma anche a un contesto semantico più ampio.

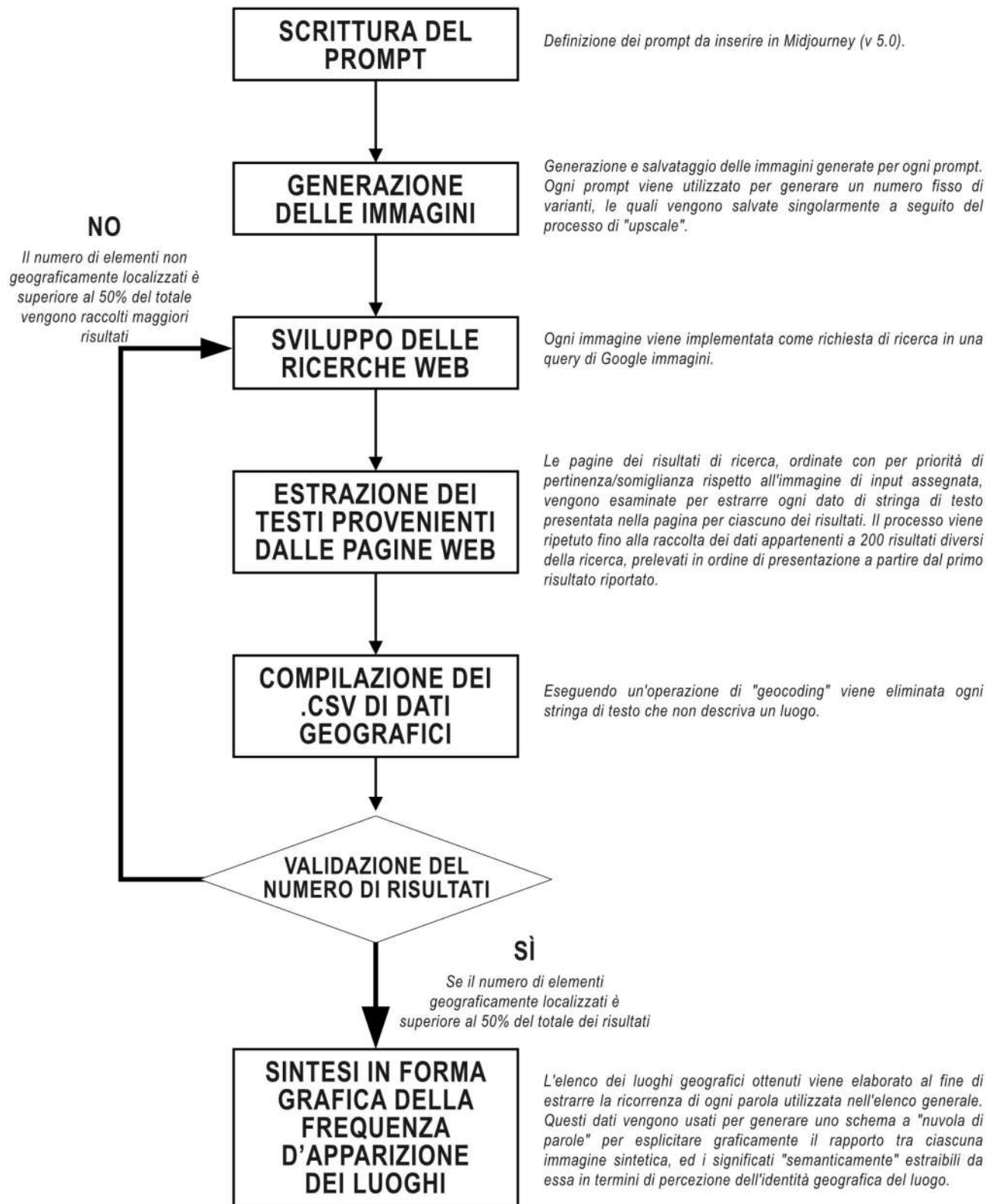
Ad esempio, le ricerche web hanno evidenziato in diverse occasioni la capacità di estrarre informazioni riguardanti aree di frontiera con un alto grado di correlazione ai soggetti indicati. Per esempio, l'immagine di un contesto urbano 'francese' è stata correttamente collegata a fonti online relative non solo al territorio della Francia, ma anche in misura minore a territori italiani come la regione del Piemonte (Collier, 1963), e in particolare la città di Torino.

La correlazione tra la Francia e Torino è interessante poiché ci troviamo in una zona di confine geografico e culturale. La vicinanza geografica e la storia condivisa hanno contribuito a una certa interazione e influenza reciproca tra le due aree. Pertanto, in linea generale non deve sorprendere che la ricerca web abbia evidenziato una correlazione tra un contesto urbano 'francese' e fonti online relative alla Francia, ma anche a territori italiani come la città di Torino. Se non fosse però che tale correlazione deriva da un'informazione 'fittizia' generata sinteticamente. Questo potrebbe suggerire che i modelli utilizzati, se adeguatamente supportati da un database sufficientemente esteso, possano rappresentare con sorprendente chiarezza la connessione geografica e culturale tra queste regioni, dimostrando le potenzialità di implementare le informazioni pertinenti (figg. 7-9).

Conclusioni

In tal senso, è possibile osservare come richieste di rappresentazioni apparentemente banali possono nascondere al contrario

Fig. 6
Workflow operativo seguito dal punto di vista metodologico per eseguire una preliminare comparazione semantica rispetto ai contenuti di rappresentazione derivati dall'uso del sistema *Midjourney* nella manipolazione di scenari urbani geograficamente categorizzati.



una complessità sottostante molto elevata. I modelli attualmente in uso nei sistemi menzionati, paiono dimostrare la capacità di collegare e rappresentare graficamente contenuti semantici con una sorprendente precisione. Mentre è necessario approfondire la dinamica di questa relazione con maggiore attenzione – il che verà approfondito in future ricerche – può essere interessante intravedere le possibilità offerte da questi strumenti non solo come mezzi di elaborazione di immagini, ma anche come strumenti di indagine e comprensione del mondo e delle relazioni umane che si sviluppano al suo interno.

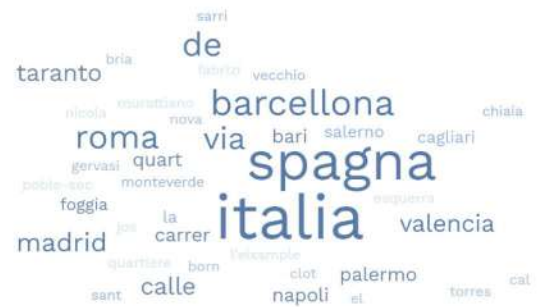
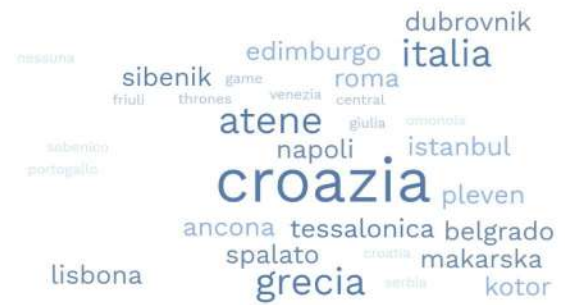
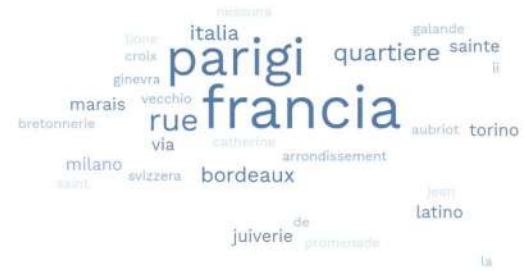
Nell'insieme, il fatto che le parole geografico-culturale (“italiano”, “francese”, etc) dei prompt utilizzati in *Midjourney* siano state ‘riconosciute’ dall’IA di *Google*, che ha suggerito risultati che a livello geografico-culturale sono strettamente legati a quelli richiesti con i prompt, ci dimostra che l’IA già oggi consente un controllo su aspetti difficilmente definibili e quantificabili, quali la cultura, il tipo di natura e lo stile architettonico, che al di là di ineliminabili bias offrono uno strumento di controllo piuttosto coerente e oggettivamente misurabile come tale.

Addirittura, gli eventuali *bias* inaspettati o ‘errori’ – quali la riconduzione di immagini di città Dalmate in passato di cultura italiana nell’ambito del ‘paesaggio italiano’ – non ci fanno tanto aborrire il *bias*, né ci spinge ad un revanscismo irredentista, ma ci interroga sul potenziale che tali sorprese possono avere nell’ottica di aiutarci a percepire collegamenti magari meno evidenti ai nostri occhi, e magari non ortodossi. Ciò, potenzialmente, ci aiuta a ‘leggere’ meglio i pezzi della realtà e a ricombinarli in un puzzle ancora meglio risolto, guidandoci a vedere relazioni tra elementi apparentemente distanti, e viceversa, un po’ come se si procedesse a ri-attribuire la paternità artistica di opere sulla base dell’implicita definizione dello ‘spirito’ o stile dell’artista così come ‘percepito’ dall’IA, potenzialmente migliorando la nostra comprensione.

Rimandiamo a successive ricerche per una più approfondita disamina dei meccanismi che portano alle sorprendenti capacità dei sistemi IA qui illustrati, e in particolare sul tema degli *embeddings* di natura semantica che – pur nella loro parziale natura di ‘*black-box*’, in cui a ciascun parametro vettoriale non si può con certezza attribuire un significato comprensibile all’uomo – possono essere comunque studiati rispetto al loro

Figg. 7-9

A sinistra dall’alto al basso: esempi di immagini elaborate tramite i *prompt* “*A photo of a beautiful French urban context*”, “*A photo of a beautiful Croatian urban context*”, “*A photo of a beautiful Spanish urban context*”. A destra: esempi di rappresentazione ‘nuvola di parole’ ottenute dall’elaborazione delle ricerche online correlate ai relativi *prompt*. La grandezza di ciascun vocabolo è direttamente proporzionale alla ricorrenza di quella stessa parola all’interno del dataset elaborato.



effetto de facto sugli esiti dei modelli di IA, e allo stesso tempo sono alla base della capacità di tali modelli di ‘cogliere’ aspetti ineffabili della realtà spesso elusivi addirittura per le persone, e che avremmo immaginato del tutto inafferrabili da parte di una ancora primitiva IA.

Bibliografia

- Alam, N. (2022). Understanding MixNMatch: Creating A More Realistic Synthetic Image. *Towardsdatascience*.<<https://towardsdatascience.com/understanding-mixnmatch-creating-a-more-realistic-synthetic-image-40f9ba039689>> (Ultimo accesso 13 maggio 2023).
- Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: a review. *Artificial Intelligence Review*, 55(1), 589-656. <https://doi.org/10.1007/s10462-021-10039-7>
- Barocas, S., & Selbst, A. D. (2016). Big Data’s Disparate Impact. *California Law Review*, 104(3), 671-732. doi: 10.2139/ssrn.2477899
- Blokpoel, M., van Kesteren, M., Stolk, A., Haselager, P., Toni, I., & van Rooij, I. (2012). Recipient Design in Human Communication: Simple Heuristics or Perspective Taking? *Frontiers in Human Neuroscience*, 6. doi: 10.3389/fnhum.2012.00253
- Bolukbasi, T., Chang, K. W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *30th Conference on Neural Information Processing Systems*.
- Bravo, L. (2010). La città storica contemporanea: genius loci e genius saeculi. *DOAJ: Directory of Open Access Journals*. doi: 10.6092/issn.2036-1602/2162
- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, 81, 1-15.
- Collier, W. (1963). French Influence on the Architecture of Filippo Juvarra. *Architectural History*, Jan. 1963. doi: 10.2307/1568282
- DALL·E 2 (s.d.).<<https://openai.com/product/dall-e-2>> (ultimo accesso 13 maggio 2023).
- de Saussure, F. (1916). *Cours De Linguistique Générale*, vol. 1. Payot.
- Dean, J. (2023). Google Research, 2022 & beyond: Language, vision and generative models. *Google Research*<<https://ai.googleblog.com/2023/01/google-research-2022-beyond-language.html#GenerativeModels>> (ultimo accesso 13 maggio 2023).

- Dhariwal, P., & Nichol, A. (2021). Diffusion Models Beat GANs on Image Synthesis. *arXiv*.<<https://arxiv.org/abs/2105.05233>> (ultimo accesso 13 maggio 2023).
- Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative Adversarial Networks, Generating 'Art' by Learning About Styles and Deviating from Style Norms. *arXiv:1706.07068*.
- Florian, M. C. (2023). Can Artificial Intelligence Systems like *DALL-E* or *Midjourney* Perform Creative Tasks? *Archdaily*.<https://www.archdaily.com/987228/can-artificial-intelligence-systems-like-dall-e-or-midjourney-perform-creative-tasks?ad_source=search&ad_medium=search_result_articles> (ultimo accesso 12 maggio 2023).
- Gazzoni, F. (1994). *Manuale di diritto privato*. Edizioni scientifiche Italiane.
- Gu, J., Wang X., Li C., Zhao C., Fu W., Liang G., & Qiu J. (2022). AI-enabled image fraud in scientific publications. *Patterns*, 3(7). doi: 10.1016/j.patter.2022.100511
- Ho, J., & Salimans, T. (2022). Classifier-Free Diffusion Guidance. *arXiv:2207.12598*.
- Inhal, (2022). I modelli di diffusione generativi in grado di superare i GAN sullasintesi delleimmagini. *Inhal.it*. <<https://ihal.it/i-modelli-di-diffusione-generativi-in-grado-di-superare-i-gan-sulla-sintesi-delle-immagini/>> (ultimo accesso 13 maggio 2023).
- Karpathy, A., Abbeel, P., Brockman, G., Chen, P., Cheung, V., Duan, Y., Goodfellow, I., Kingma, D., Ho, J., Houthoof, R., Salimans, T., Schulman, J., Sutskever, I., & Zaremba, W. (2016). Generative models. *Open AI*.<<https://openai.com/research/generative-models>> (ultimo accesso 13 maggio 2023).
- Laaziri, M., Benmoussa, K., Khouliji, M., Larbi, K. M., & Yamami, A. E. (2019). Analyzing bootstrap and foundation font-end frameworks: a comparative study. *International Journal of Electrical and Computer Engineering*, 9(1). doi: 10.11591/ijece.v9i1.pp713-722
- McClard, P. (2022). Sorry, Generating AI Art Does Not Make You an Artist. *Medium*.<<https://medium.com/mlearning-ai/sorry-generating-ai-art-does-not-make-you-an-artist-108fe16d4de7>> (ultimo accesso 13 maggio 2023).
- Midjourney*, (s.d.). <<https://www.midjourney.com/app/>> (ultimo accesso 13 maggio 2023).
- Muppalla, V., & Hendryx, H. (2023). Diffusion Models: A Practical Guide. *Scale*.<<https://scale.com/guides/diffusion-models-guide#diffusion-models-why-are-they-important>> (ultimo accesso 12 maggio 2023).
- Nichol, A., Dhariwal, P., Ramesh, A., Shyam, P., Mishkin, P., McGrew, B., Sutskever, I., & Chen, M. (2021). GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models. *arXiv:2112.10741*

- Peirce, C. S. (1954). *Collected Papers of Charles Sanders Peirce*. Harvard University Press.
- Photography (1855). *The Crayon*, 1(11), 170.
- Photon* (2023). Civitai. <<https://civitai.com/models/84728/photon>> (ultimo accesso 15 febbraio 2024).
- Reilly, B. F. (1979). The Early Work of John Moran, Landscape Photographer. *American Art Journal*, 11(1), 65. doi: 10.2307/1594133
- Shapley, G. (2011). After the artefact: Post-digital photography in our post-media era. *J Vis Art Pract*, 10(1), (5-20). doi: 10.1386/jvap.10.1.5_1
- Shen, C., Kasra, M., Pan, W., Bassett, G. A., Malloch, Y., & O'Brien, J. F., (2019). Fake images: The effects of source, intermediary, and digital media literacy on contextual assessment of image credibility online. *New Media Soc*, 21(2), 438-463. doi: 10.1177/1461444818799526
- Singh, P. N., & Gowdar, T. P. (2021). Reverse Image Search Improved by Deep Learning. *2021 IEEE Mysore Sub Section International Conference (MysuruCon)*, 596-600. <https://doi.org/10.1109/MysuruCon52639.2021.9641572>
- Stable Diffusion*, (s.d.). <<https://stablediffusionweb.com/>> (ultimo accesso 13 maggio 2023).
- Teicher, J. G. (2016). When Photography Wasn't Art. *Jstor Daily*.<<https://daily.jstor.org/when-photography-was-not-art/>> (ultimo accesso 13 maggio 2023).
- The Ultimate Prompting Guide or How to Talk with an AI Being a Dummy (2022). *Prompthero*.<<https://prompthero.com/stable-diffusion-prompt-guide>> (ultimo accesso 13 maggio 2023).
- Wankhede, C. (2023). What is Midjourney AI and how does it work? How does a state-of-the-art AI image generator work? We break it down. *Android authority*. <<https://www.androidauthority.com/what-is-midjourney-3324590/>> (ultimo accesso 12 maggio 2023).
- Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 9(11), 39-52. doi: 10.22215/timreview/1282
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. *arXiv*. doi: 10.48550/arxiv.2302.11382
- Witteveen, S., & Andrews, M. (2022). Investigating Prompt Engineering in Diffusion Models. *arXiv*: 2211.15462
- Zemel, R., Wu, Y., Swersky, K., Pitassi, T., & Dwork, C. (2013). Learning fair representations. *ICML'13: Proceedings of the 30th International Conference on International Conference on Machine Learning*, 28(3), 325-333.

Fonti delle immagini

Fig. 1. Elaborazione grafica degli autori realizzata utilizzando *Midjourney*, versione 5.0, 2023 (<https://www.midjourney.com>), *prompt* “A photo of a beautiful riverbank in an urban context”.

Fig. 2. Elaborazione grafica degli autori realizzata utilizzando (Stable Diffusion, s.d.) Pre-trained model *Photon* (2023).

Fig. 3. Elaborazione grafica degli autori realizzata utilizzando *Midjourney* (s.d.).

Fig. 4. Elaborazione grafica degli autori.

Fig. 5. Elaborazione grafica degli autori realizzata utilizzando (Stable Diffusion, s.d.) Pre-trained model *Photon* (2023).

Fig. 6. Elaborazione grafica degli autori.

Figg. 7-9. Elaborazioni grafiche degli autori realizzate utilizzando *Midjourney* (s.d.).